

# Evaluation of NU-WRF rainfall forecasts for IFloodS

Di Wu<sup>1,2</sup>, Christa Peters-Lidard<sup>3</sup>, Wei-Kuo Tao<sup>1</sup>, and Walter Petersen<sup>4</sup>

*1 Mesoscale Atmospheric Processes Laboratory  
NASA Goddard Space Flight Center  
Greenbelt, Maryland*

*2 Science Systems and Applications, Inc.  
Lanham, Maryland*

*3 Hydrological Sciences Laboratory  
NASA Goddard Space Flight Center  
Greenbelt, Maryland*

*4 Code 610.W  
NASA GSFC/Wallops Flight Center  
Wallops Island, Virginia*

-----

<sup>1</sup> Corresponding author address: Di Wu, Code 612, NASA Goddard Space Flight Center, Greenbelt, MD 20771, Email: di.wu@nasa.gov

## Abstract

The Iowa Flood Studies (IFloodS) campaign was conducted in eastern Iowa as a pre-GPM-launch campaign from 1 May to 15 June 2013. During the campaign period, real time forecasts are conducted utilizing NASA-Unified Weather Research and Forecasting (NU-WRF) model to support the everyday weather briefing. In this study, two sets of the NU-WRF rainfall forecasts are evaluated with Stage IV and Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation (QPE), with the objective to understand the impact of Land Surface initialization on the predicted precipitation. NU-WRF is also compared with North American Mesoscale Forecast System (NAM) 12 km forecast. In general, NU-WRF did a good job at capturing individual precipitation events. NU-WRF is also able to replicate a better rainfall spatial distribution compare with NAM. Further sensitivity tests show that the high-resolution makes a positive impact on rainfall forecast. The two sets of NU-WRF simulations produce very close rainfall characteristics. The Land surface initialization do not show significant impact on short-term rainfall forecast, and it is largely due to the soil conditions during the field campaign period.

51

## 52 **1. Introduction**

53 One of the goals of the Global Precipitation Measurement (GPM) mission  
54 ground validation program is to conduct integrated hydrological validation, in which  
55 the terrestrial water budget is utilized to evaluate the accuracy of blended satellite  
56 and/or model-based precipitation products. The Iowa Flood Studies (IFloodS)  
57 campaign was conducted in eastern Iowa as a pre-GPM-launch campaign from 1 May  
58 to 15 June, 2013, with the goal of examining how well GPM and other blended  
59 products could be used for flood forecasting.

60 The areas of focus for the IFloodS campaign were the Cedar and Iowa River  
61 Basins, which were covered by a ground-based NASA NPOL radar deployed along  
62 with rain gauges and disdrometers in addition to the existing NEXRAD radar network  
63 (Cunha et al., 2015). In addition to the traditional rainfall-oriented instrumentation,  
64 in the South Fork Iowa River, a network of rain gauge and soil moisture platforms  
65 was deployed in coordination with the Agricultural Research Service and NASA's  
66 Soil Moisture Active Passive (SMAP) mission (Coopersmith et al, 2015). These soil  
67 moisture platforms, in addition to extensive existing streamflow monitoring by the  
68 USGS and Iowa Flood Center, help support the integrated hydrologic validation goals  
69 of the campaign. Further, additional, high-resolution integrated hydrologic validation  
70 is supported in the Turkey River Basin, within which 20 rain gauges with soil  
71 moisture probes and two Iowa Flood Center XPOL weather radars were deployed.

72 To support deployment of ground-based instrumentation, our team at GSFC  
73 conducted real-time forecasting with a meteorological model, which was delivered

daily to support 0900LT forecast briefings delivered to the campaign personnel. This effort required not only dedicated computational resources, but also a robust modeling system capable of simulating severe convective episodes typical of eastern Iowa during the active spring period. In this work we will provide a comprehensive evaluation of the modeling system, including the precipitation forecasts from two different configurations designed to evaluate the impact of land surface initialization on the two forecasts. We first describe the experimental design, including the modeling system, configuration and evaluation datasets. Next, we present an evaluation of the precipitation forecasts based on an archive for the entire experimental period relative to ground data in addition to an operational forecast model. Finally, we discuss the implications of this work for future forecasting applications.

## **1. Experiment design**

### **a. NU-WRF model**

The NASA-Unified WRF (NU-WRF; <http://nuwrf.gsfc.nasa.gov>) modeling system has been developed at Goddard Space Flight Center (GSFC) as an observation-driven integrated modeling system that represents aerosol, cloud, precipitation and land processes at satellite-resolved scales (Peters-Lidard et al. 2015). NU-WRF is a superset of the National Center for Atmospheric Research (NCAR) Advanced Research WRF (ARW) dynamical core model, achieved by fully integrating the GSFC Land Information System (LIS; Kumar et al. 2006; Peters-Lidard et al. 2015), the WRF/Chem enabled version of the GOddard Chemistry Aerosols Radiation Transport (GOCART; Chin et al. 2000) model, the Goddard Satellite Data Simulation Unit (G-SDSU; Matsui et al. 2009),

and custom boundary/initial condition preprocessors. Several NASA physical packages (microphysics and radiation) have also been implemented into NU-WRF. These physical processes include CRM-based microphysics (Tao et al. 2003; Lang et al. 2007, 2011, 2014) and radiation (Chou and Suarez 1999) schemes. All the above features are combined into a single software release, with source code available by agreement with NASA/GSFC.

In this study, NU-WRF version 3.4.1 (based on NCAR WRF-ARW version 3.4.1) is employed to conduct high-resolution simulations. There are 60 vertical levels and 3 spatial domains with 9, 3 and 1 km grid spacing (Fig. 1), and time steps of 27, 9 and 3 seconds respectively. The Grell-Devenyi cumulus parameterization scheme (Grell and Devenyi 2002) is adopted for the outer domain; no convective parameterization was used for two inner domains. The PBL parameterization employed the Mellor-Yamada-Janjic (Mellor and Yamada 1982) Level-2 turbulence closure model through the full range of atmospheric turbulent regimes. The Goddard broadband two-stream (upward and downward fluxes) approach was used for the short- and long-wave radiative flux calculations (Chou and Suarez 1999) and its explicit interactions with clouds (microphysics). In addition, the numerical simulations use Goddard 3ICE scheme (Lang et al. 2011), which prognoses three types of ice hydrometeor species (i.e. cloud ice, snow, and graupel).

#### **b. NASA Land Information System**

The Land Information System (LIS) is a core component of NU-WRF. It is a flexible land surface modeling and data assimilation framework developed with the goal

of integrating satellite- and ground-based observational data products and advanced land surface modeling techniques to produce optimal fields of land surface states and physics (Kumar et al. 2006; Peters-Lidard et al. 2007). The infrastructure can not only be directly coupled with the atmosphere, it can also integrate high resolution observations with the model forecasts to generate improved estimates of land surface conditions such as soil moisture, evaporation, snow pack, and runoff, at 1km and finer spatial resolutions and at one-hour and finer temporal resolutions.

The role of LIS in the simulation was two-fold: First, to provide physically consistent land surface initialization for NU-WRF; Second, to interact with the surface layer and atmospheric components of NU-WRF and produce coupled water, energy and momentum fluxes. The LSM employed in LIS for this study is Noah LSM version 3.2.1 (Ek et al. 2003). WRF-ARW version 3.4.1 uses Noah LSM version 3.4.1. For consistency, the Noah land surface model is run offline within LIS using the same domain configuration as NU-WRF. The Noah LSM in the offline LIS also uses the same soil and vegetation database as NU-WRF.

The offline LIS run is cold started from 2 May 2008 to 1 May 2013. The long spin up period is used for land surface states to achieve equilibrium for initialization of WRF-LIS. The LIS offline spin up uses Stage IV rainfall data to provide hourly rainfall, and utilizes GDAS to provide atmospheric forcing input. The Stage IV rainfall product is only used to provide forcing for the LSM during the analysis cycle and is not assimilated into the atmospheric component of the coupled simulation.

#### **c. NU-WRF real-time forecasts**

During the IFloodS campaign period, two sets of 48-hour NU-WRF forecasts were produced twice a day from May 1<sup>st</sup> to June 15<sup>th</sup> 2013. These forecasts required 7 hours to produce with 2048 CPUs on NASA NCCS supercomputer.

The configuration for both the control (WRF) and coupled (COUP) simulation are as described above (in section 3a). The most substantial difference between the two sets of forecasts are the initialization of the soil states in the land surface model. As previously mentioned, the control simulation uses the same version of Noah Land Surface model as in WRF V3.4.1, while the coupled simulation (COUP) uses a slightly older version (3.2.1) of the Noah Land-Surface model, since that is the most recent version of Noah that was implemented in LIS at the time of the campaign. The changes from 3.2 to 3.4 focused on snow, and based on other offline analyses (not shown) did not result in significant differences in soil moisture, runoff, or land surface fluxes. Hence, the key difference between the WRF and COUP runs is the soil initialization. The soil initialization for the control simulations comes from spatially interpolating the soil moisture and soil temperature states in the forcing dataset, which is NAM in our case, while the land-surface initialization for COUP is produced from a three-year (May 1<sup>st</sup> 2010 to May 1<sup>st</sup> 2013) offline spin up by LIS, updated daily by an analysis forced by a previous WRF forecast where the precipitation is bias-corrected using the Stage IV blended radar-rain gauge product.

As shown in Fig. 2, the forecast starts everyday at 00 UTC and 12 UTC for 48 hours integration. For coupled simulation, every forecast cycle starts with a short LIS offline analysis, and then proceeds to an online coupling between WRF and LIS. The offline analysis is initialized from the previous day NU-WRF forecast and forward

integrated for 24 hours to current initialization time. Stage IV data was used to provide hourly precipitation forcing for the LIS analysis, while NU-WRF output from previous day supplements the atmospheric forcing. Similar to the three-year offline spin up, the analysis provides the soil initialization for the coupled simulation, while the atmospheric initialization is provided from NAM, as in the control (WRF) simulation.

#### **d. Data and methodology**

Two observational datasets are employed for model evaluation. The Stage IV (Lin and Mitchell 2005) rainfall estimates are available at 4 km grid spacing every hour. It is a widely used rainfall product by both hydrological and meteorological communities, due to its national coverage, high spatial and temporal resolutions and overall low biases (e.g. Tang et al. 2014, Wu et al. 2011, and Seo et al. 2013). Its good performance in Mean Squared Error (MSE) and total bias contributes from the effectiveness of bias correction, and the manual Quality Control (QC) procedures (Cunha et al. 2015).

Multi-Radar/Multi-Sensor System (MRMS) Quantitative Precipitation Estimate (QPE) integrates radar QPE, gauge QPE, local gauge bias corrected radar QPE, and gauge orographic precipitation climatology QPE. MRMS QPE has a 2 min time interval in each  $0.01^\circ \times 0.01^\circ$  grid box. One improvement to the previous QPE product ‘Q2’ is that MRMS uses most advanced dual-polarimetric (DP) radar technologies to eliminate non-meteorological echoes, and it also provides a more accurate spatial distribution of precipitation.

Despite the advantages in DP QPE in certain aspects, it does not necessarily provide an overall superior QPE than single-polarimetric (SP) QPE, such as Stage IV. According to Cunha et al. (2015), DP QPE shows a higher MSE than Stage IV estimates.



However, Stage IV also shows a decreased correlation with rain gauges with increasing rainfall threshold (greater than 5 mm/h) than DP estimates. In this study, we choose MRMS as a QPE reference for most statistics analysis. Stage IV is also shown to give a uncertainty range between two datasets.

The model analysis is performed on the inner most grid of NU-WRF forecasts, which output each hour with 1 km grid spacing. NAM forecasts are available at three hourly intervals with 12 km spatial resolution. Both 00 Z and 12 Z forecasts are evaluated for 48 hours model integration on each day for all modeling datasets during May 1<sup>st</sup> to June 15 2015. All datasets are remapped onto the NAM grid and are intercompared at three hourly time intervals.

### **3. Evaluation of rainfall forecast**

Figure 3 shows the accumulated precipitation from a composite of NU-WRF simulations and Stage IV dataset from May 1<sup>st</sup> to June 15<sup>th</sup> 2013. Stage IV shows clearly a relatively higher accumulated rainfall area from Iowa to northern Illinois compare to surrounding regions. NU-WRF simulated rainfall accumulation shows a similar spatial pattern as the Stage IV. The high accumulation is over Iowa and Northern Illinois, with a peak that has a higher magnitude than Stage IV accumulated rainfall.

#### **a. Rainfall time-series**

Figure 4 shows three hours accumulated rainfall from two NU-WRF simulations (WRF and COUP), NAM, Stage IV, and MRMS from May 1<sup>st</sup> to June 15<sup>th</sup>. The values for models are the mean from available forecast cycles. The first 6 hours simulations are

considered as model spin up period, thus are removed from the analysis. Despite the overall overestimation comparing with observations, NU-WRF runs capture the individual precipitation events well. There are only slight differences between the two sets of NU-WRF runs. NAM is very close to observation for averaged rainfall over the whole period. However, it does not replicate the individual events well, especially during May 16<sup>th</sup> to June 1<sup>st</sup>, where NAM tends to miss the peak or significantly underestimates the rainfall. During the rest of the periods, NAM has the tendency for overestimation. So despite NAM has a better averaged rainfall over the six-week period than NU-WRF, the NAM does not show a good forecast skill for individual events.

Also shown in Fig. 4, the precipitation events come in groups. There are seven wet periods can be identified (Table 1), grouped by at least one dry day (or very light rain) in between two periods. Each precipitation period can be caused by single convective/precipitating system or may be a succession of convective systems. Most of these events have strong upper level support, which is typical for spring and early summer events. Three of these periods have either short wave troughs (VII) or a combination of short wave and long wave troughs (III and IV), which brings weak but complicated forcings to the region. Especially for period IV, where one short wave comes after another, and it associates with a series of propagating systems at surface.

NAM struggles to produce an accurate precipitation forecast during period IV (Fig. 4). On the other hand, NU-WRF is able to reproduce the individual peaks during period IV. Despite using NAM as initial and boundary condition, the finer resolution NU-WRF simulations certainly show improved features in reproducing these precipitation events. These series of weakly forced events demonstrate the characteristics

of warm season MCSs. And according to many studies (e.g. Doswell et al. 1996; Fritsch and Carbone 2004), it is very difficult to improve forecasts for deterministic warm season rainfall events. Even with the advancements in numerical models in recent years in both physics and resolutions, the forecast for warm season convections still remain a challenge. And since it brings most heavy rainfall and events with high societal impacts, it is also quite important to improve warm season deterministic rainfall forecast.

#### **b. Rainfall statistics**

The domain averaged three hours accumulated rainfall statistical scores are calculated from May 1<sup>st</sup> to June 15<sup>th</sup> with respect to the forecast lead-time (Fig. 5). The bias scores (Fig. 5a) are negative for WRF, COUP, and NAM in the first six hours, which is due to the cold start of these models, where all the precipitation values are initialized from zero. The two NU-WRF simulations have appreciably higher biases scores than NAM comparing with MRMS, which is also evident from Fig. 4 that NAM underestimates rainfall during period III and IV while overestimates for other periods. Despite the lower bias, NAM has higher RMSE (0.73) than two NU-WRF simulations (0.67 and 0.69). Also NAM has lower correlation (0.60) comparing with two NUM-WRF simulations (0.76 and 0.73). Both NU-WRF domain averaged rainfall has relatively high correlations with MRMS, which echoes that NU-WRF has captured individual precipitation events in Fig. 4. The correlation trend between models and MRMS are only decreasing slightly with increase of forecast lead-time. The correlation between Stage IV and MRMS are very high (0.98). The differences between two NU-WRF simulations are small enough comparing with the differences between Stage IV and

MRMS. However, the differences between two NU-WRF correlations increase with forecast lead-time.

When considering spatial variability between different datasets and MRMS, the correlation scores are much lower than the area averaged quantities (Fig. 6). The correlation between Stage IV and MRMS decreases from 0.98 for area-averaged rainfall to 0.82 for considering both time and spatial correlations. The score drops significantly for correlations between models and MRMS, from previously 0.7 to under 0.2. From Fig. 6, the decreasing trend of correlation is obvious for models, which demonstrate the forecast skill has decreased with increase of forecast hours. Despite the low spatial correlation, NU-WRF seems to produce consistently slightly higher correlation than NAM forecasts.

Table 2 shows the spatial correlation scores for all seven different forecast periods. NAM shows consistently lower scores than both NU-WRF forecasts. And the correlation scores vary with different periods. NAM has the lowest correlation during period IV and VII, while NU-WRF has a relatively lower correlation during V and VII. Period IV is one difficult period for NAM, on the other hand, NU-WRF did a fine job at capturing individual events (not shown) during this period. It is one of the periods that NAM and NU-WRF has the largest differences in correlation score (0.07 vs 0.18).

One caveat of spatial correlations analysis is that the displacement in spatial correlations is heavily penalized, same as the appearance of spurious precipitating regions. Thus for individual time slice, the correlation may not necessarily reflect all aspects of the forecast performance. Such as in Fig. 7, even the NU-WRF has captured the characteristics of heavy precipitation, it still has a lower correlation score. However,

with a relatively bigger sample size, the correlation is capable of describing a general trend. Such as, NAM consistently have a lower spatial correlation score than NU-WRF (Fig. 6), which is consistent with that NAM has a lower domain averaged correlation as well (Fig. 5c). On the other hand, identification for spatial displacement is also important. It is the dominant source of quantitative precipitation forecast (QPF) error (Ebert and McBride 2000). Poor QPF skill has hindered hydrologic applications, particularly streamflow forecasting operations (Cuo et al. 2011).

The two NU-WRF runs have very close correlations with each other, and the differences grow with time, which is also observed from rainfall spatial distributions for various cases (not shown). It is a question whether the spread is caused by physical differences between the two models or by random error growth. From Fig. 6a, the differences are fluctuating after 12 hours into the forecast for all cases, and similarly for different periods. However, a systematic evaluation of this error behavior is beyond the scope of this study.

Figure 8 shows rainfall statistics for each forecast cycle using MRMS as a reference. So the value at May 15<sup>th</sup> shows correlation of MRMS and 6 to 48 hours model forecast initialized at 00 UTC on May 15<sup>th</sup>. The overall scores are similar to the scores from Fig. 5. The differences are due to the sample selection and whether or not including the first six hours into the considerations. In addition to aforementioned differences between models, it also clearly shows how models perform during different periods. Fig. 8a shows models underestimate some rainfall events during period III, IV, and VII, while overestimate the rest of the periods. NAM shows large negative bias during period IV and positive bias during other periods, so the overall low bias of NAM is merely an

averaged effect from positive and negative biased cases. For the RMSE score, all models show similar score ( $\sim 0.55$ ) compare with MRMS. The spread of the RMSE scores among models (0.02) are even smaller than that between Stage IV and MRMS (0.14). High model RMSE are also associated with heavy precipitation periods. Period III and IV produce many heavy precipitation events, also associated with high model RMSEs. Figure 8c shows the correlations between different datasets and MRMS. NAM performed poorly during period III and IV, which is also reflected in Fig. 4. The low correlation period is around May 11<sup>th</sup> to 13<sup>th</sup>, where not much rainfall was brought to the area (Fig. 4). Small phase shift in rainfall measurements can induce low correlation scores between Stage IV and MRMS. Models also have poor correlations with MRMS during this period. The differences of correlation scores of the two NU-WRF simulations (0.01) are smaller than that of Stage IV and MRMS (0.05).

### **c. Rainfall time-series PDF**

The precipitation PDF is also evaluated for model and observational datasets. As shown on Fig. 9, NU-WRF compares very well with Stage IV data for rainfall accumulations less than 8 mm every three hours, but there are some overestimations for heavier rainfall frequencies. On the other hand, NAM has produced outstandingly high frequencies for very light precipitation (0.25-0.5 mm), while significantly underestimated the frequencies for heavy rainfall ( $> 4$  mm). This feature is also evident from rainfall spatial distribution (Fig. 7), NAM produces large light rain area, but fails to produce any heavy rainfall. It is also quite common in many coarse resolution models; large grid spacing limits its ability to resolve sub-grid convections.

The two sets of NU-WRF runs have very close PDF distribution. The differences of the two NU-WRF runs are even smaller than those between the two observational dataset (Stage IV and MRMS). MRMS has higher light rainfall (0.25-1 mm) coverage and higher coverage for intense rainfall (>16 mm). The small differences between two NU-WRF runs indicate that different land surface initialization and differences within the LSM do not have a big impact on precipitation intensity. On the other hand, the PDF distribution is shown to be more sensitive to different cloud microphysics schemes (Tao et al., 2015).

## **4 Discussion**

### **a. Sensitivity Tests**

NU-WRF and NAM share the same initial and boundary conditions, but their predicted rainfall characteristics are very different (shown in section 3), especially during period IV. Additional sensitivity tests are performed for 9 km (WRF\_9km) and 3km (WRF\_3km) resolution, using 1 km NU-WRF without LIS coupling (WRF) as control run. WRF\_9km employs only the outer most domain in Fig. 1, while the WRF\_3km uses two outer domains in Fig.1. Besides with different resolutions and domain setups, WRF\_9km uses Grell-Devenyi ensemble cumulus scheme (GD), where the WRF and WRF\_3 km runs only applied GD for the outer most grid (with 9 km resolution). NAM has significantly underestimated rainfall and the peak is out of phase with the observations for forecast initiated at 00 UCT on May 29<sup>th</sup> 2013 (Fig. 10). WRF and COUP, on the other hand, captured the peak, despite 3 hours delayed than the observation. By reducing the resolution from original 1 km (WRF) to 9 km (WRF\_9km),

there is a reduction of peak value and shift of phase. However, the change of resolution from 1 km to 3 km (WRF\_3km) does not result in significant changes in the forecast. In addition, 9 km run with Betts-Miller-Janjic (BMJ) cumulus scheme (WRF\_BMJ) is used to compare with 9 km with the GD cumulus scheme (WRF\_9km), same GD scheme was applied to the outer most domain in the control run (WRF). BMJ cumulus scheme is also the one used in NAM forecast. WRF\_BMJ has an even lower peak magnitude than WRF\_9km run, and its averaged rainfall is the most comparable to NAM in all the simulations. There is improved forecast skill by adopting finer resolution and using GD cumulus scheme, even when initial and boundary conditions stay the same. This is also consistent with many previous studies (Wang and Seaman 1997; Gallus 1999), the choice of convective schemes have a strong influence on simulated rainfall pattern.

#### **b. Soil moisture and rainfall**

Of particular interest in this study is that whether there are improvements in rainfall forecast by applying high resolution and more accurate land-surface initial condition comparing to interpolated fields from regional model forecast. As shown in previous sections, the differences in the rainfall forecasts between the two NU-WRF runs are rather small. One possible reason for such small differences is that the region of interest is under the influence of many heavy precipitation events during the campaign period, thus with high water availability, the moisture transport from surface are similarly high in both models. One indication for the above argument is high evaporative fraction (EF), which is the ratio of latent heat to available energy at the land surface. EF is a diagnostic for the surface energy balance (energy-limited state or a moisture-limited



state), supposedly isolates soil moisture and vegetation from radiation and turbulent factors. Despite the strong diurnal periodicity for two components in the surface energy balance, EF is generally considered to be a constant during daytime hours (Nichols and Cuenca 1993; Crago 1996; Crago and Brutsaert 1996). Figure 11 shows the daily EF from NU-WRF run, which is averaged from local 7 am to 6 pm. EF stays over 0.6 for most days, which means the energy fluxes to the surface energy budget are mainly contributed by latent heating. With high EF, the impact for precipitation processes with different soil initialization is minimized. Figure 12 shows the evaporative fraction at local noon for NU-WRF forecast initialized at 12 UTC on May 30<sup>th</sup> 2013, and it also shows the 0-10 cm soil moisture at model initialization. WRF has higher soil moisture than COUP (Figure 12 and 13), but COUP has an only slightly smaller EF than WRF. The slightly dryer top soil moisture from COUP is also observed in Georgia and South Carolina during summer season, which is actually closer to the observed soil moisture from the U.S. Department of Agriculture's Soil Climate Analysis Network (SCAN) (Case et al. 2011). The small EF difference shows that their similar partition in surface energy budget. WRF also has a much lower resolution than COUP. The lower resolution is due to the interpolated soil moisture field from NAM that has a 12 km resolution, while COUP uses LIS offline spin up that provides soil moisture at a resolution that is consistent with NU-WRF grid. In addition, COUP uses Stage IV observed rainfall to force the LIS offline spin up, which should result a more accurate and observational consistent soil moisture profile than WRF. Despite with fine resolution and more accurate soil moisture initialization, the high EF indicates COUP and WRF have a similar

land surface moisture transport to atmospheric boundary layer, which will minimize their impact on precipitation processes.

## **5. Conclusions**

Two sets of NU-WRF are used for providing real-time forecasts twice a day for two-day long integration during IFloodS field campaign from May 1<sup>st</sup> to June 15<sup>th</sup>. One of the NU-WRF forecast uses NAM interpolated land surface field as LSM forcing; the other one uses LIS spin up to provide land surface conditions, which assimilates the latest Stage IV observed precipitation. These two sets of model datasets are compared with low resolution forcing dataset (NAM) and with each other. The precipitation fields are evaluated with Stage IV and MRMS. Two observations datasets are able to provide a difference range, which indicates observational uncertainties. The main conclusions are as follows:

- 1) Both NU-WRF simulations are able to reproduce the individual precipitation event during the field campaign period, which NAM is out of sync with observations for heavy precipitation events during period IV. In addition, for those events where rainfall intensity less than 1 mm/h, NAM tends to overestimates the rainfall amount. However, for heavy rainfall events (e.g. May 20<sup>th</sup>, May 25<sup>th</sup> to May 30<sup>th</sup>, and June 13<sup>th</sup>), NAM has underestimated the rainfall amounts. Despite NAM has a better averaged rainfall over the six-week period compared with NU-WRF, the NAM forecast skill is not necessarily better for individual events. Even though NU-WRF shares the same initial and boundary condition from NAM, and its rainfall field is averaged onto the coarser NAM grid, NU-WRF still outperforms NAM in both time and spatial correlations.

2) NU-WRF is also able to produce a better rainfall PDF than NAM. NAM significantly underestimates the frequencies for heavy rainfall and largely overestimates frequencies for very light rainfall. While NU-WRF is able to produce PDF that is very close to the observed distribution from Stage IV.

3) NU-WRF sensitivity tests show that by switching to a coarser resolution and to a different convective scheme, the rainfall forecast skill has reduced, which turns out to be more comparable to NAM forecast in a case study. Model resolution makes a difference for rainfall forecast, but it also depends on which scale it applies to, whether within the convective permitting scale or to coarser scales.

4) LIS spin up with Stage IV forcing has the advantage of producing higher resolution and more accurate surface properties than without LIS spin up. However, the benefit for the precipitation forecast is marginal. Two sets of NU-WRF simulations do not yield significant differences on rainfall characteristics during IFloodS field campaign period. The differences between two NU-WRFs are much smaller than the differences between NAM and NU-WRF or between the two observational datasets (Stage IV and MRMS). Evaporative fraction indicates the relatively similar land surface moisture transport between the two NU-WRF simulations, which inhibit the land-surface improvement to have a positive impact on precipitation forecast.

#### **Acknowledgements**

This research was supported by the NASA Precipitation Measurement Missions (PMM) and Modeling and Analysis Program (MAP) solicitations through awards to PI's Peters-Lidard, Tao and Petersen. This support is gratefully acknowledged. We also

442 acknowledge the excellent computational and storage support provided by NASA's  
443 Center for Climate Simulation (NCCS).

444

## References

- Carbone, R. E., J. D. Tuttle, D. A. Ahijevych, and S. B. Trier, 2002: Inferences of Predictability Associated with Warm Season Precipitation Episodes. *J. Atmos. Sci.*, **59**, 13.
- Case, J. L., S. V. Kumar, J. Srikishen, and G. J. Jedlovec, 2011: Improving Numerical Weather Predictions of Summertime Precipitation over the Southeastern United States through a High-Resolution Initialization of the Surface State. *Wea. Forecasting*, **26**, 785–807. doi:<http://dx.doi.org/10.1175/2011WAF2222455.1>
- Chin, M., R. B. Rood, S.-J. Lin, J. F. Muller, and A. M. Thomspon, 2000: Atmospheric sulfur cycle in the global model GOCART: Model description and global properties, *J. Geophys. Res.*, **105**, 24,671–24,687.
- Chou, M.-D., and M. J. Suarez, 1999: A shortwave radiation parameterization for atmospheric studies, *NASA/TM-104606*, **15**, pp 40.
- Coopersmith, Evan J., Michael H. Cosh, Walt A. Petersen, John Prueger, and James J. Niemeier, 2015: Soil Moisture Model Calibration and Validation: An ARS Watershed on the South Fork Iowa River. *J. Hydrometeor.*, **16**, 1087–1101. doi: <http://dx.doi.org/10.1175/JHM-D-14-0145.1>.
- Crago, R., 1996: Conservation and variability of the evaporative fraction during the daytime. *J. Hydrol.*, **180**, 173–194.
- Crago, R., and W. Brutsaert, 1996: Daytime evaporation and the self-preservation of the evaporative fraction and the Bowen ratio. *J. Hydrol.*, **178**, 241–255.

467 Cunha, Luciana K., James A. Smith, Witold F. Krajewski, Mary Lynn Baeck, Bong-Chul  
 468 Seo, 2015: NEXRAD NWS Polarimetric Precipitation Product Evaluation for  
 469 IFloodS. *J. Hydrometeor.*, Early online release, doi:  
 470 <http://dx.doi.org/10.1175/JHM-D-14-0148.1>  
 471 Cuo, L., T. C. Pagano, and Q. J. Wang, 2011: A review of quantitative precipitation  
 472 forecasts and their use in short- to medium-range streamflow forecasting. *J.*  
 473 *hydrometeor.*, **12**, 713-729  
 474 Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.  
 475 D. Tarpley, 2003: Implementation of Noah land surface model advances in the  
 476 National Centers for Environmental Prediction operational mesoscale Eta model,  
 477 *J. Geophys. Res.*, **108**(D22), 8851, doi:10.1029/2002JD003296  
 478 Gallus, W. A., Jr., 1999: Eta simulations of three extreme rainfall events: Impact of  
 479 resolution and choice of convective scheme. *Wea. Forecasting*, **14**, 405–426.  
 480 Gallus, W. A., and M. Segal, 2000: Sensitivity of forecast rainfall in a Texas convective  
 481 system to soil moisture and convective scheme. *Wea. Forecasting*, **15**, 509–526.  
 482 Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection  
 483 combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**,  
 484 Article 1693.  
 485 Kumar, S. V., Y. Tian, C. Peters-Lidard, and Coauthors, 2006: Land information system:  
 486 An interoperable framework for high resolution land surface modeling. *Environ.*  
 487 *Modelling Software*, **21**, 1402-1415.

488 Lang, S., W.-K. Tao, R. Cifelli, W. Olson, J. Halverson, S. Rutledge, and J. Simpson,  
 489 2007: Improving simulations of convective system from TRMM LBA: Easterly  
 490 and Westerly regimes, *J. Atmos. Sci.*, **64**, 1141-1164.

491 Lang, S. E., W.-K. Tao, X. Zeng, and Y. Li, 2011: Reducing the biases in simulated radar  
 492 reflectivities from a bulk microphysics scheme: Tropical convective systems, *J.*  
 493 *Atmos. Sci.*, **68**, 2306–2320.

494 Lang, S., W.-K. Tao, J.-D. Chern, D. Wu, and X. Li, 2014: Benefits of a 4th ice class in  
 495 the simulated radar reflectivities of convective systems using a bulk microphysics  
 496 scheme, *J. Atmos. Sci.*, **71**, 3583-3612.

497 Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses:  
 498 Development and application. Preprints, *19th Conf. on Hydrology, Am. Meteorol.*  
 499 *Soc.*, San Diego, CA, P1-2.

500 Matsui, T., X. Zeng, W.-K. Tao, H. Masunaga, W. S. Olson, and S. Lang, 2009:  
 501 Evaluation of long-term cloud-resolving model simulations using satellite  
 502 radiance observations and multi-frequency satellite simulators. *J. Atmos. Oce.*  
 503 *Tech.*, **26**, 1261-1274.

504 Matsui, T., J. Santanello, J. J. Shi, W.-K. Tao, D. Wu, C. Peters-Lidard, E. Kemp, M.  
 505 Chin, D. Starr, M. Sekiguchi, and F. Aires, 2014: Introducing Multi-Sensor  
 506 Satellite Radiance-based Evaluation for Regional Earth System Modeling, *J.*  
 507 *Geophys. Res.*, **119**, 8450-8475, doi: <http://dx.doi.org/10.1002/2013JD021424>

508 Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for  
 509 geophysical fluid problems, *Rev. Geophys. Space Phys.*, **20**, 851-875.

510 Nichols, W. E., and R. H. Cuenca, 1993: Evaluation of the evaporative fraction for  
 511 parameterization of the surface energy balance. *Water Resour. Res.*, **29**, 3681–  
 512 3690.

513 Peters-Lidard, C.D., E. M. Kemp, T. Matsui, J. A. Santanello, Jr., S. V., Kumar, J. P.  
 514 Jacob, T. Clune, W.-K. Tao, M. Chin, A. Hou, J. L. Case, D. Kim, K.-M. Kim,  
 515 W. Lau, Y. Liu, J.-J. Shi, D. Starr, Q. Tan, Z. Tao, B. F. Zaitchik, B. Zavodsky,  
 516 S. Q. Zhang, M. Zupanski (2015), Integrated modeling of aerosol, cloud,  
 517 precipitation and land processes at satellite-resolved scales, *Environmental*  
 518 *Modelling & Software*, **67**, 149–159.  
 519 doi:<http://dx.doi.org/10.1016/j.envsoft.2015.01.007>

520 Seo, B.-C., L. K. Cunha, and W. F. Krajewski, 2013: Uncertainty in radar- rainfall  
 521 composite and its impact on hydrologic prediction for the eastern Iowa flood of  
 522 2008. *Water Resour. Res.*, **49**, 2747–2764.

523 Tang, L., Y. Tian, and X. Lin, 2014: Validation of precipitation retrievals over land from  
 524 satellite-based passive microwave sensors. *J. Geophys. Res. Atmos.* **119** (8): 4546-  
 525 4567 [10.1002/2013JD020933]

526 Tao, W.-K., J. Simpson, D. Baker, S. Braun, M.-D. Chou, B. Ferrier, D. Johnson, A.  
 527 Khain, S. Lang, B. Lynn, C.-L. Shie, D. Starr, C.-H. Sui, Y. Wang and P. Wetzell,  
 528 2003: Microphysics, radiation and surface processes in the Goddard Cumulus  
 529 Ensemble (GCE) model, A Special Issue on Non-hydrostatic Mesoscale  
 530 Modeling, *Meteor. Atmos. Phys.*, **82**, 97-137.

531 Tao, W. K., D. Wu, S. Lang, J. Chern, C. Peters-Lidard, A. Fridlind, and T. Matsui, 2015:  
 532 High-resolution NU-WRF simulations of a deep convective-precipitation system



533 during MC3E: Part I: Comparisons between Goddard microphysics schemes and  
 534 observations. *J. Geophys. Res. Atmos.*, in revision.

535 Wang, W., and N. L. Seaman, 1997: A comparison study of convective schemes in a  
 536 mesoscale model. *Mon. Wea. Rev.*, **125**, 252–278.

537 Wu, W., D. Kitzmiller, and S. Wu, 2011: Evaluation of radar precipitation estimates from  
 538 the National Mosaic and Quantitative Precipitation Estimation System and the  
 539 WSR-88D Precipitation Processing System over the Conterminous United States.  
 540 *J. Hydrometeor.*, **13**, 1080–1093.

541

Period	Dates	Synoptic
I	5/2-5/5	Trough and surface front
II	5/8-5/11	Low
III	5/16-5/24	Short wave trough followed by long wave trough
IV	5/25-6/2	Group of short wave trough
V	6/4-6/7	Low
VI	6/8-6/11	Trough
VII	6/12-6/14	Two short wave troughs

542 Table 1. Seven precipitation periods and their synoptic setups.

543  
 544  
 545

Period	I	II	III	IV	V	VI	VII
WRF	0.22	0.20	0.15	0.18	0.11	0.16	0.09
COUP	0.22	0.21	0.15	0.18	0.11	0.16	0.12
NAM	0.20	0.17	0.10	0.07	0.09	0.13	0.07

Table 2. Rainfall spatial correlations between model forecasts with MRMS during the seven precipitation periods.

## **List of Figures**

Figure 1: NU-WRF grid configuration. The outer domain (labeled 1 at the center) has a horizontal resolution of 9 km. The middle domain (labeled 2) has a horizontal resolution of 3 km, and the inner domain (labeled 3) has a horizontal resolution of 1 km and covers Iowa.

Figure 2: Flow chart for real-time forecast using NU-WRF coupling with LIS. Orange boxes show the initial time of forecast cycles. Green boxes show LIS offline spin up to provide soil initial condition. The LIS spin up uses Stage IV to provide precipitation forcing and NU-WRF previous day forecast to provide additional meteorological forcing. Blue boxes show NU-WRF coupling with LIS uses LIS offline spin up to provide surface initial conditions and use NAM to provide meteorological initial conditions and boundary conditions.

Figure 3: Accumulated precipitation from NU-WRF real-time forecast and Stage IV dataset from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

Figure 4: Three hours accumulated precipitation from NU-WRF with LIS (CP) and without LIS coupling (WRF), NAM, MRMS, and Stage IV datasets from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

Figure 5: Time series of bias (a), root mean square error (b), and correlation (c) of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against MRMS with respect to forecast hours.

Figure 6: Similar to Figure 5 (c), but for spatial correlations between NU-WRF and NAM forecasts with MRMS for the whole campaign period from May 1<sup>st</sup> to June 15<sup>th</sup> (a), for period III May 16<sup>th</sup> to 24<sup>th</sup>(b), and for period IV May 25<sup>th</sup> to June 2<sup>nd</sup> (c). Stage IV has a 0.82 correlation averaged through out the campaign period.

Figure 7: Three hours rainfall accumulation (mm) for MRMS, NAM, NU-WRF with LIS coupling (COUP) and without LIS (WRF) at 06 UTC on May 20<sup>th</sup> 2013, which is 18 hours since model initialization.

Figure 8: Time series of bias (a), root mean square error (RMSE) (b), and correlation (c) of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against MRMS with respect to each forecast.

Figure 9: PDF of observed and forecasted three hours accumulated precipitation from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

Figure 10: Domain averaged three hours rainfall accumulation for NU-WRF sensitivity runs. The result is from forecast cycle initiated at 00 UTC on May 29<sup>th</sup>, 2013.

Figure 11: Daily averaged (7 am to 6 pm) evaporative fraction from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

Figure 12: Evaporative fraction at local noon (17 UTC on May 19<sup>th</sup>) and top soil moisture (0-10 cm) at model initialization, which is at 12 UTC on May 19<sup>th</sup>, 2013.

Figure 13: Same as Figure 11, except for daily averaged soil moisture.

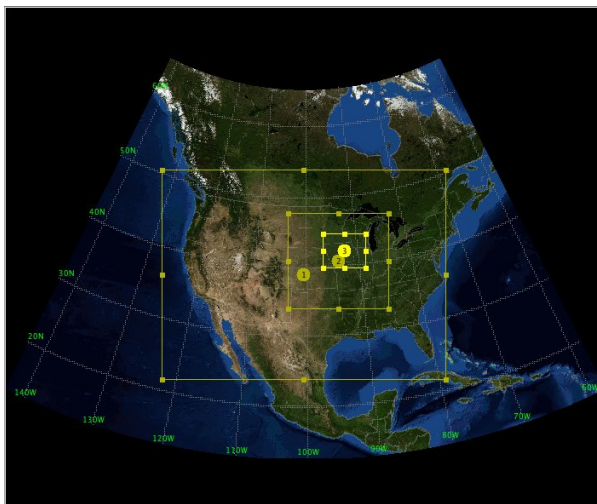


Figure 1: NU-WRF grid configuration. The outer domain (labeled 1 at the center) has a horizontal resolution of 9 km. The middle domain (labeled 2) has a horizontal resolution of 3 km, and the inner domain (labeled 3) has a horizontal resolution of 1 km and covers Iowa.

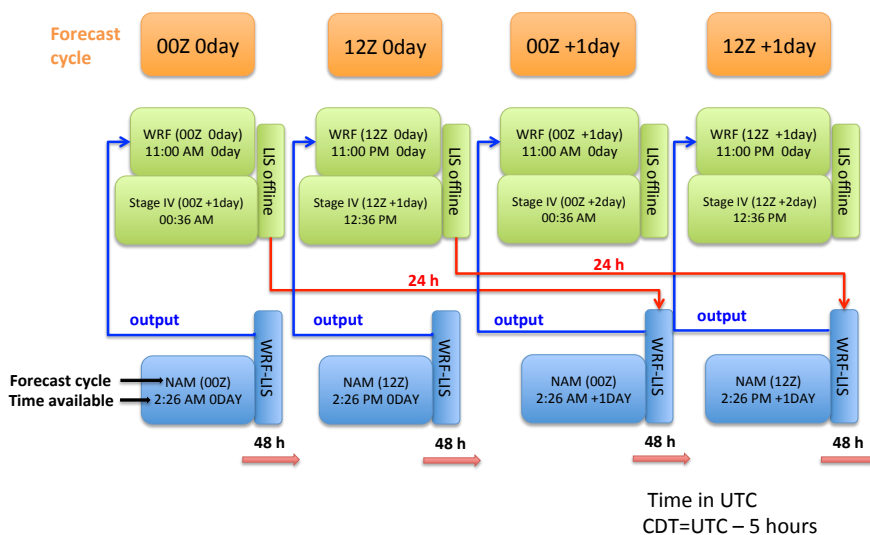


Figure 2: Flow chart for real-time forecast using NU-WRF coupling with LIS. Orange boxes show the initial time of forecast cycles. Green boxes show LIS offline spin up to

provide soil initial condition. The LIS spin up uses Stage IV to provide precipitation forcing and NU-WRF previous day forecast to provide additional meteorological forcing. Blue boxes show NU-WRF coupling with LIS uses LIS offline spin up to provide surface initial conditions and use NAM to provide meteorological initial conditions and boundary conditions.

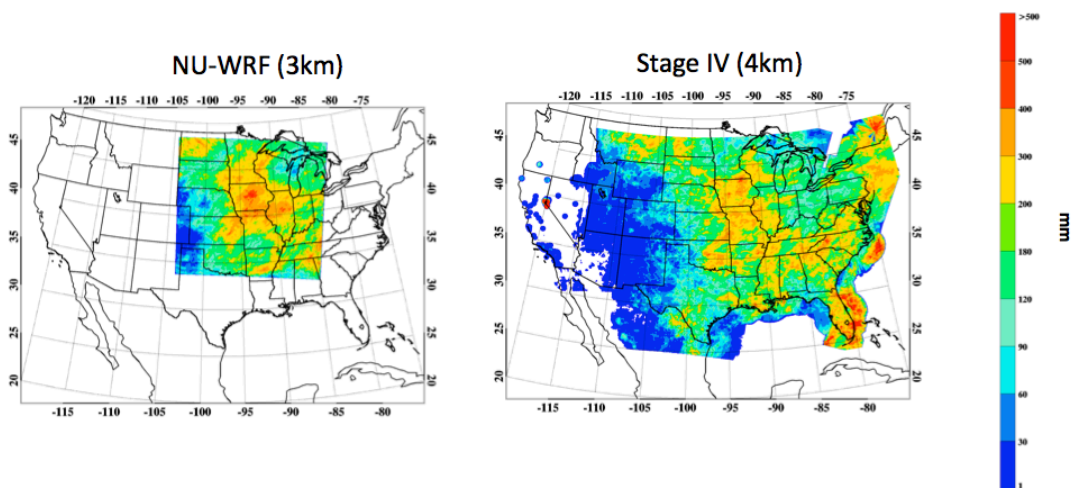


Figure 3: Accumulated precipitation from NU-WRF real-time forecast and Stage IV dataset from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

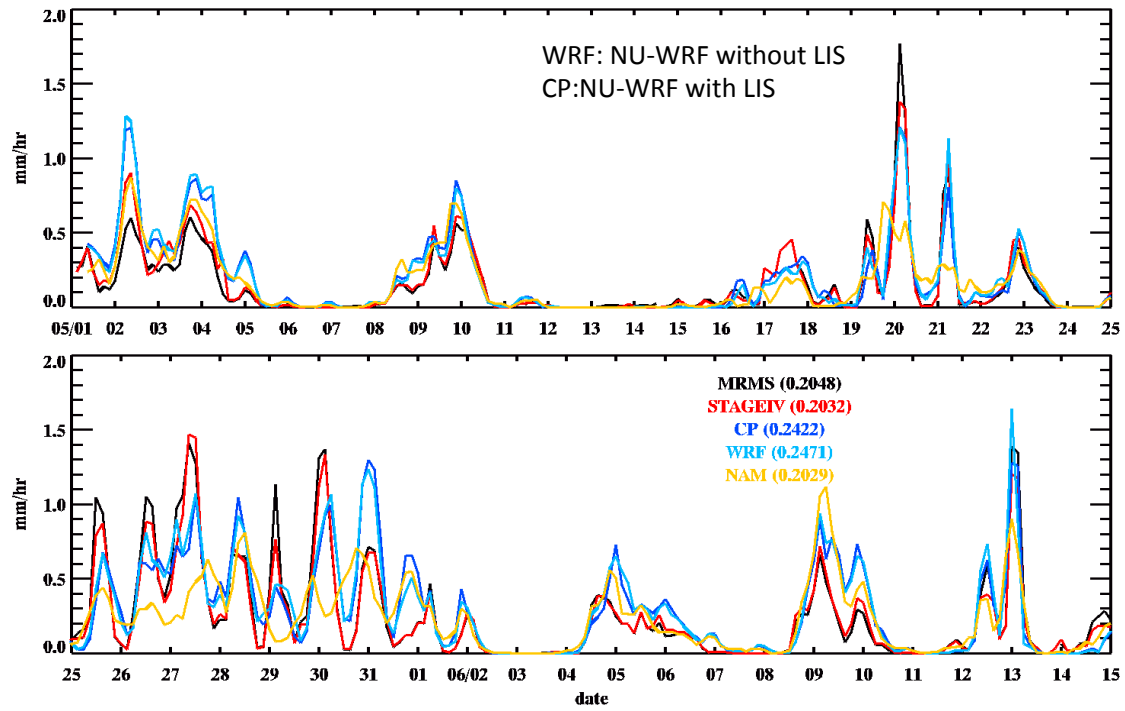


Figure 4: Three hours accumulated precipitation from NU-WRF with LIS (CP) and without LIS coupling (WRF), NAM, MRMS, and Stage IV datasets from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

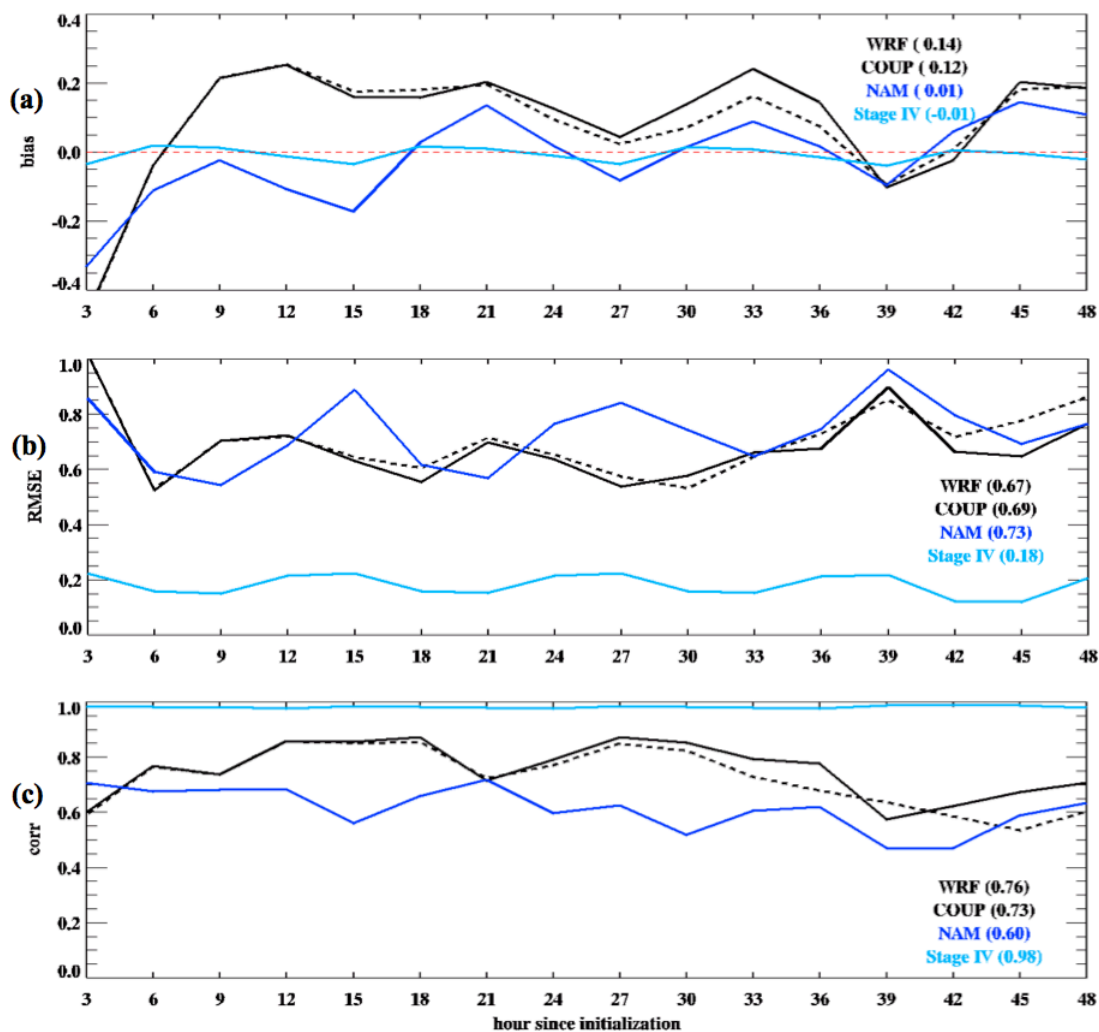


Figure 5: Time series of bias (a), root mean square error (b), and correlation (c) of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against MRMS with respect to forecast hours.

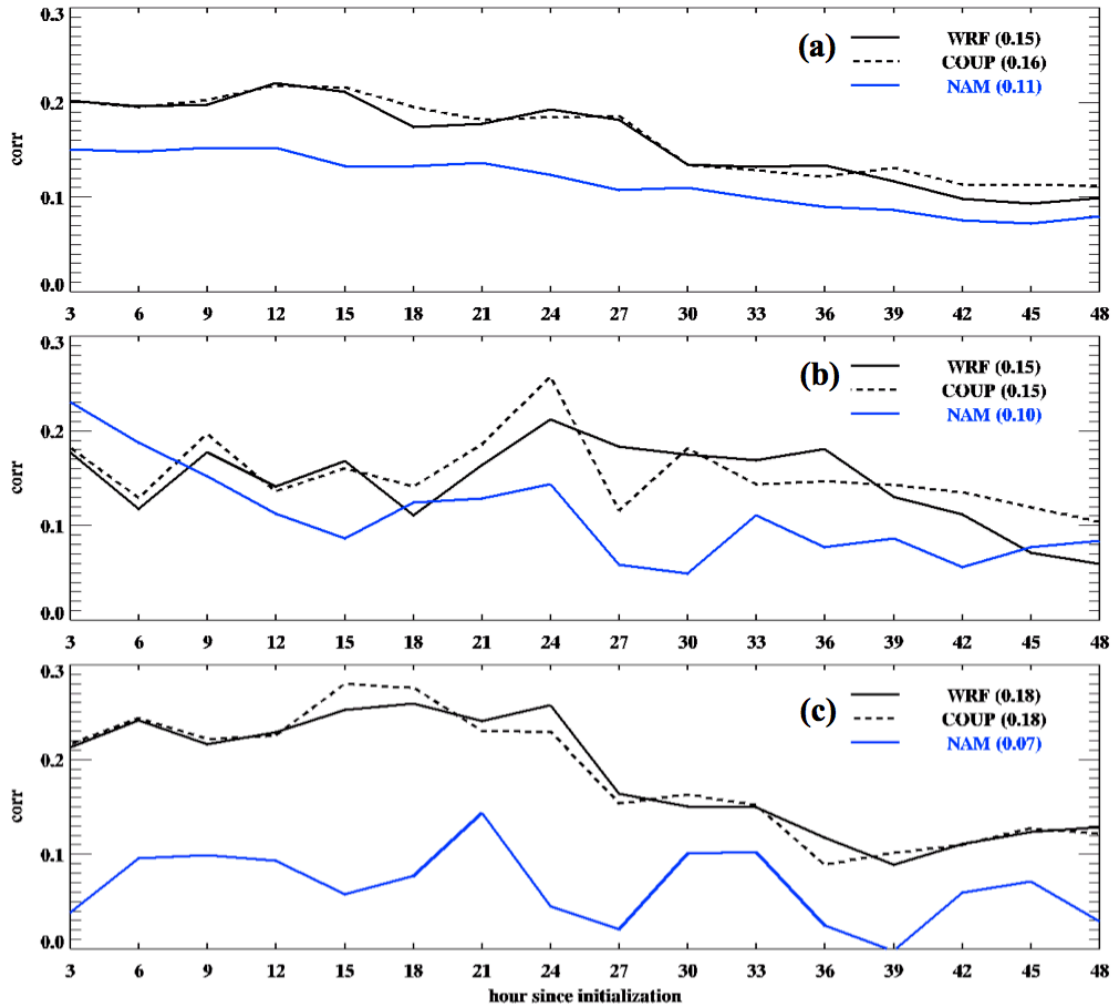
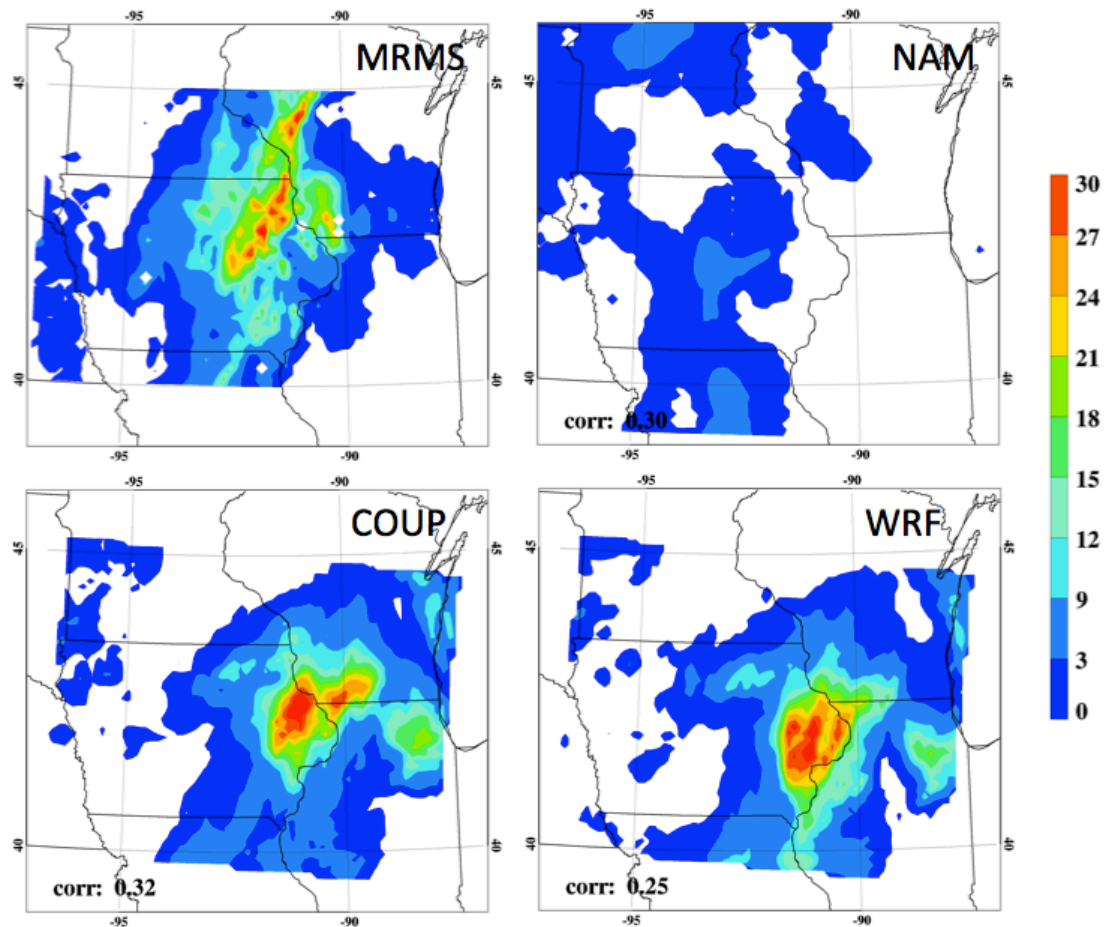


Figure 6: Similar to Figure 5 (c), but for spatial correlations between NU-WRF and NAM forecasts with MRMS for the whole campaign period from May 1<sup>st</sup> to June 15<sup>th</sup> (a), for period III May 16<sup>th</sup> to 24<sup>th</sup> (b), and for period IV May 25<sup>th</sup> to June 2<sup>nd</sup> (c). Stage IV has a 0.82 correlation averaged through out the campaign period.





630

631 Figure 7: Three hours rainfall accumulation (mm) for MRMS, NAM, NU-WRF with LIS

632 coupling (COUP) and without LIS (WRF) at 06 UTC on May 20<sup>th</sup> 2013, which is 18

633 hours since model initialization.

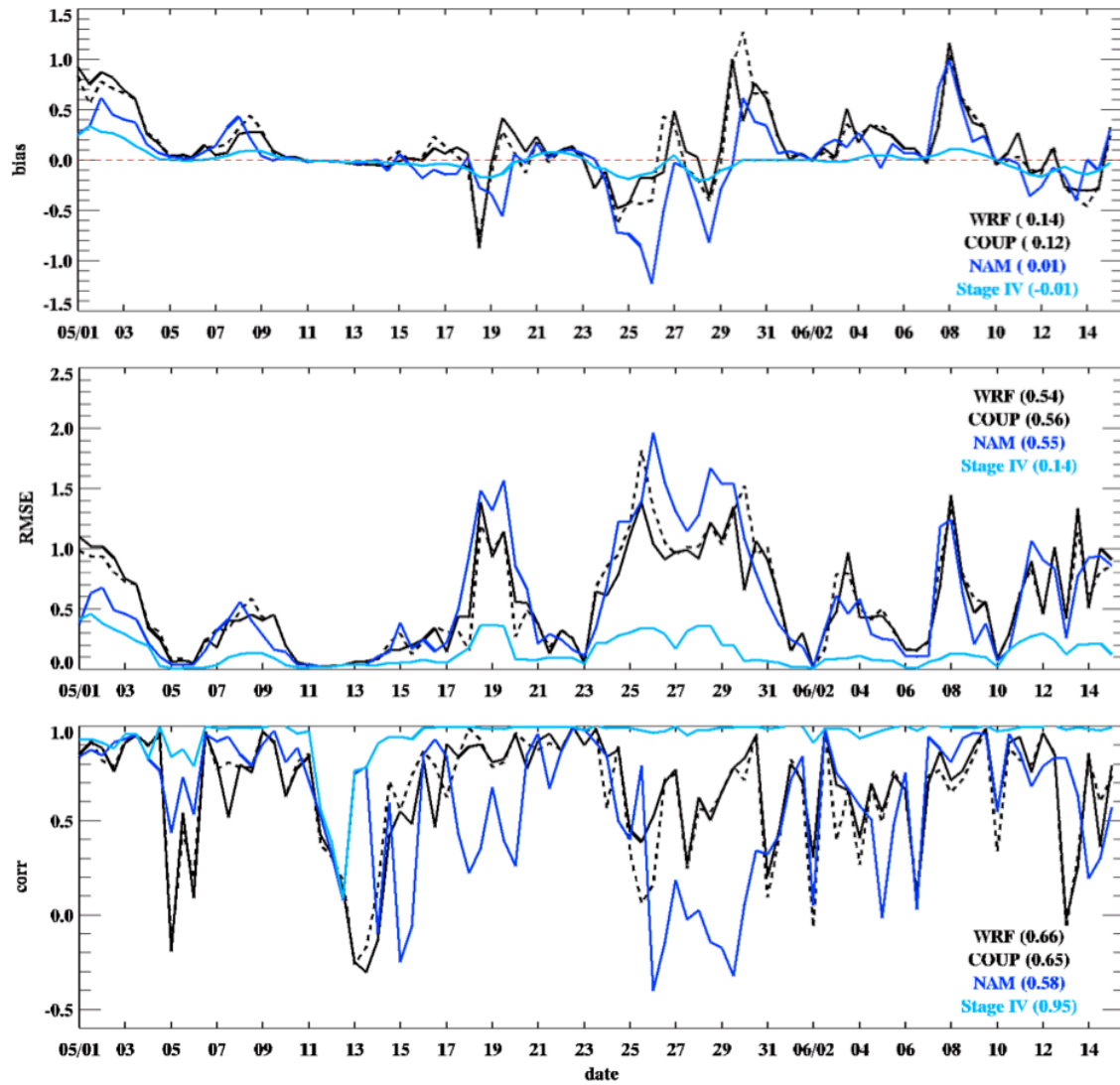


Figure 8: Time series of bias (a), root mean square error (RMSE) (b), and correlation (c) of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against MRMS with respect to each forecast.

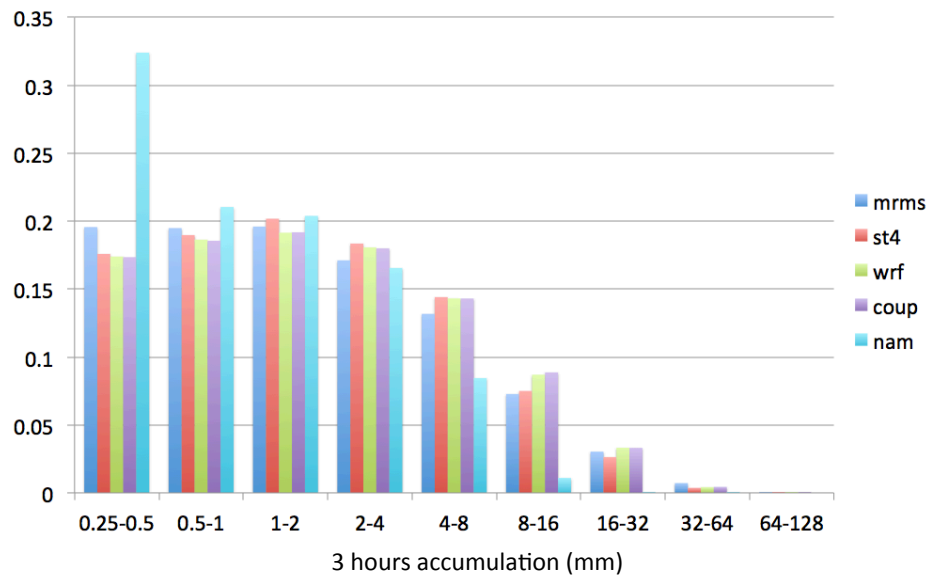


Figure 9: PDF of observed and forecasted three hours accumulated precipitation from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

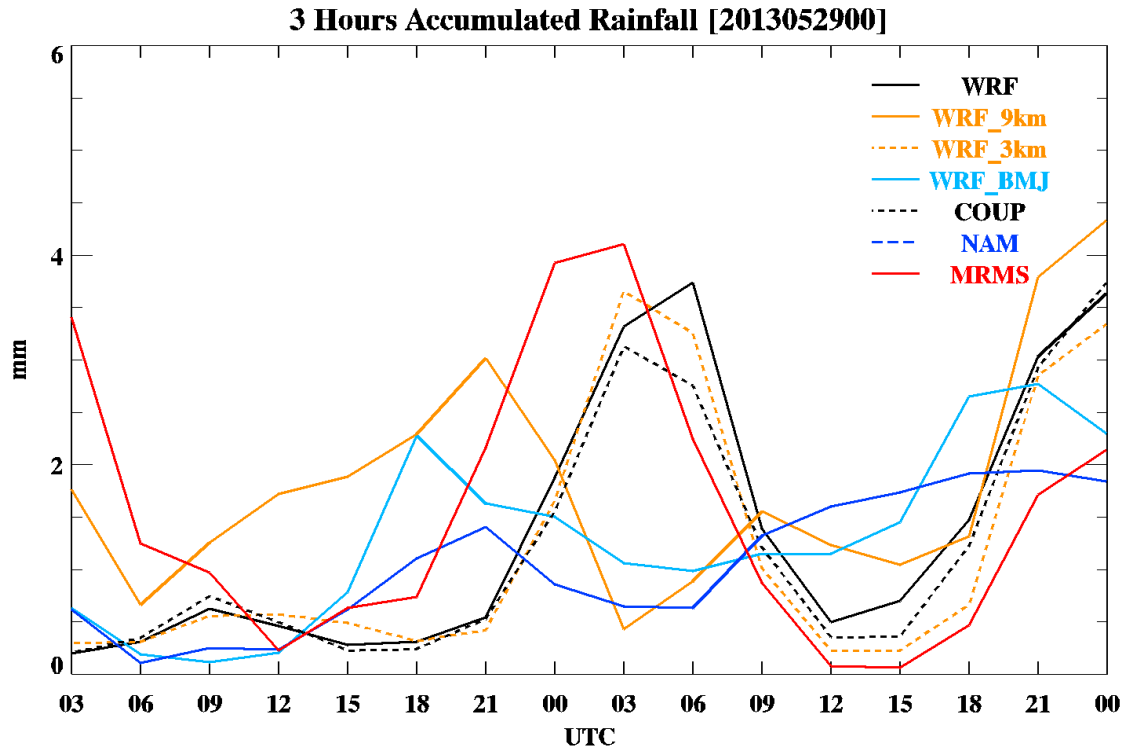


Figure 10: Domain averaged three hours rainfall accumulation for NU-WRF sensitivity runs. The result is from forecast cycle initiated at 00 UTC on May 29<sup>th</sup>, 2013.

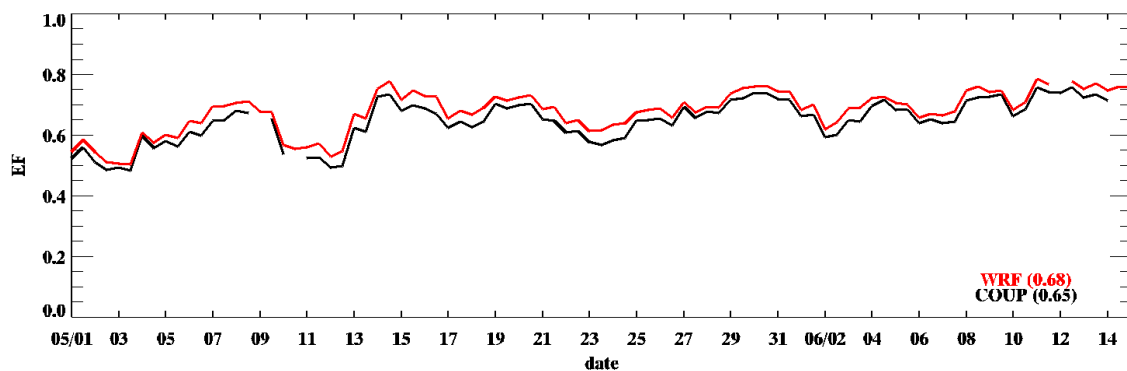


Figure 11: Daily averaged (7 am to 6 pm) evaporative fraction from May 1<sup>st</sup> to June 15<sup>th</sup> 2013.

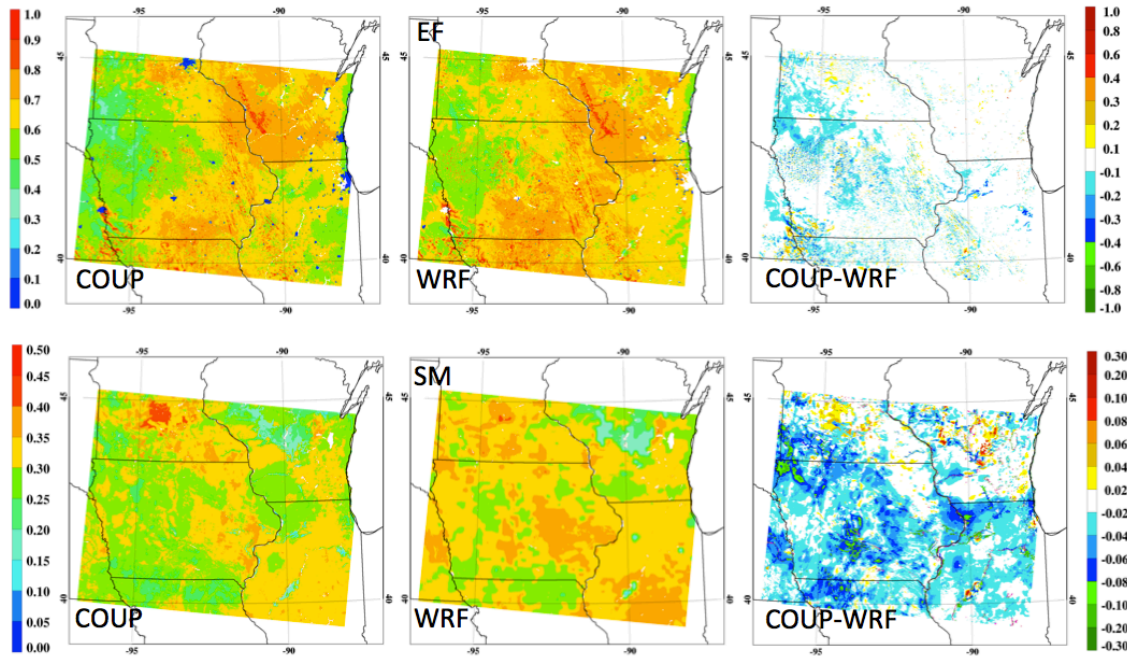


Figure 12: Evaporative fraction at local noon (17 UTC on May 19<sup>th</sup>) and top soil moisture (0-10 cm) at model initialization, which is at 12 UTC on May 19<sup>th</sup>, 2013.

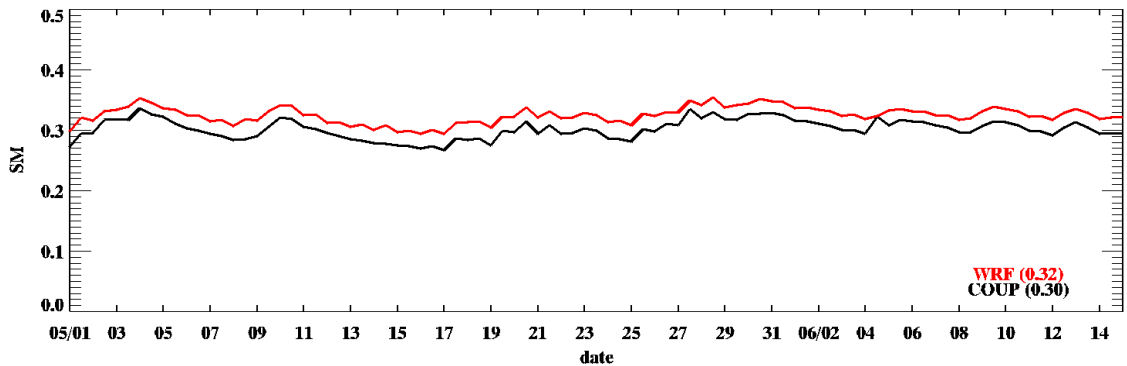


Figure 13: Same as Figure 11, except for daily averaged top soil moisture (0-10 cm).